Valuing Preservation

Abby Smith

Abstract

Preservation has value to society over and above the value of the content that is preserved. It is important to articulate this value in order to argue compellingly for the creation of public policies and economic models that adequately support preservation of culturally significant content. This article explores the societal value of preservation, discussing why questions about societal value arise in the context of the explosive growth of digital information and why they are qualitatively different questions from the ones that arose when the world knew only analog communication technologies. It assesses various ways to think about the value that inheres in content, particularly the distinctive attributes of cultural content that have societal value. It identifies benefits that preservation as such brings to society, over and above the sum of the value of the content preserved. It also examines the range of public policy issues that arise in light of the social values identified, virtually none of which are currently protected by law or regulation. In light of these societal values, it argues that the preservation community needs to collaborate with other sectors crucially dependent on long-term access to significant content to develop strategies that: make it easier and cheaper to preserve content; provide incentives and rewards for individuals and organizations to preserve; and protect the public interest in privately held content.

Introduction

As a society, we in the United States value access to information. In our roles both as citizens and as consumers, we prize our national tradition of public libraries, government archives, free press, and now, seemingly un-

LIBRARY TRENDS, Vol. 56, No. 1, Summer 2007 ("Preserving Cultural Heritage," edited by Michèle V. Cloonan and Ross Harvey), pp. 4–25

hindered access to online information via the World Wide Web. Current public policy debates surrounding access to information are concerned primarily with ensuring an equitable and just allocation of the costs and benefits of access. However, most people—even those leading the debates about copyright, licensing, and open access—are thinking about access today. They are not aware of the economic costs, or societal benefits, of providing access to cultural content over an extended period of time. This lack of understanding on the part of the general public and many public policy advocates has proven to be a significant stumbling block in securing adequate resources to preserve our analog collections. It has emerged as potentially an even bigger impediment in our endeavors to retool our preservation infrastructures to assure long-term access to digital content, including digital cultural heritage. In recent years preservation professionals have taken to characterizing their work as provision of "persistent access," "life-cycle management of information assets," "sustainability," or "stewardship" in the hope of underscoring the societal value of preservation. Is anyone listening? Does the public care? Will they be willing to pay the price for preservation, however we call it, in the twenty-first century?

While the scope and quantity of resources necessary for analog-based content preservation are not known in vivid quantitative detail by professionals, there has emerged in the past half century a common understanding about what primary analog preservation cost factors are, how to build economies of scale into their provision, and how to restructure organizations and create inter-institutional collaborations to afford them.[1] The widespread adoption of digital technology for creation of and access to culturally significant content is scarcely a decade old, if one takes online sharing of digital resources as the tipping point for adoption. There are nonetheless a number of efforts vigorously underway to understand the economic impacts of managing digital content for long-term access. How much we can afford to collect and preserve; how much we can afford to lose; and how much all of this will cost in human and financial terms—answers to these questions are critical for making the preservation case to funders. They are also crucially important, largely ignored, public policy matters.

Many professionals suspect that preservation of digital content is even more resource-intensive than that of analog, if only because there is more content, used by more people, to capture and preserve (Lyman & Varian, 2003). How high the cost will be a shock, no matter what kinds of value we may point to as the result of public investment to ensure that the digital record of today be available next week, next year, for the next generation and the next after them. Without clarifying to the public why it is important to society in general and to individuals in particular to make long-term commitments of resources to the collection and preservation of cultural content, it is unlikely to happen. And without such an understanding, we will not be able to make judicious and equitable decisions

about how those costs should be allocated among the various private and public sector constituencies.

This essay is intended to frame the question of the societal value of preservation within the context of contemporary U.S. society and public policy. A salient feature of our culture is the degree to which we extend analog and digital communication technologies into all aspects of civic and private life and have become, for all intents and purposes, critically dependent on these technologies to live safe, productive, and meaningful lives. Failure to nurture a stable and reliable information environment will put a good deal of our well-being and safety at risk. I wish to foreground here the priority claimed by preserved cultural content, that which has embedded in it historical experience and meaning that are constitutive elements of that stability and reliability. The integrity and historical continuity of cultural content are a matter of the highest priority to our society, for reasons I shall argue below. Moreover, the preservation of that cultural content rightly should be viewed as a matter of public trust, something that transcends individual or particular interests or ownership and that demands public resources and public policies to protect it. I shall use the term *cultural content* in a completely non-technical way, in the hope that it will be understood in an inclusive sense whose full detonations will become clear as we proceed. Culture can be understood as "the totality of socially transmitted behavior patterns, arts, beliefs, institutions, and all other products," more specifically "intellectual and artistic activity and the works produced by it" (*American Heritage Dictionary*, 2000, p. 442). I shall refer to the recorded works produced by culture as *content*.[2]

I shall begin by discussing why questions about societal value arise in the context of the explosive growth of digital information and why they are qualitatively different questions from the ones that arose when the world knew only analog communication technologies. Then I will assess various ways to think about the value that inheres in content, particularly the distinctive attributes of cultural content that have societal value. I will identify benefits that preservation as such brings to society, over and above the sum of the value of the content preserved. I will close by suggesting the range of public policy issues that arise in light of the social values identified, virtually none of which are currently protected by law or regulation—a curious position for a public trust. While begging the ultimate question of cost allocations—who should pay—I will argue for why we, as a society, should be willing to pay.

What's the Problem?

Why does the exponentially expanding scale of digital content creation present a novel challenge to society's willingness to pay for preservation? After all, information explosions are nothing new. They are an inevitable consequence of any innovation in recording media. Looking back

only 150 years, we see the production of inexpensive wood pulp paper, the development of audio and visual recording media, the invention of magnetic tape—all these engineering and manufacturing feats resulted in a boom in content production and a subsequent boom in content consumption following along within decades. Each in turn challenged traditional practices of stewardship, both technically and conceptually. Following this historic pattern, digital communication technologies have certainly accelerated the demand for access to information. They have given rise to an explosion of professional, amateur, and "pro/am" audiences for and auteurs of content of all kinds, including digital representations of and information about cultural content. Content owners and distributors have expanding audiences for vintage recorded sound and moving image, from Duke Ellington and Glenn Gould reissues to early Alfred Hitchcock films and the original *Twilight Zone*. Collecting institutions of all stripes, from the Metropolitan Museum of Art and the Hermitage Museum to the Library of Congress and the "Google Five" (now Six) research libraries have invested significant resources to extend the reach of their artifactual collections through digitization.[3] Content distribution companies and memory institutions conspire to feed—and in turn create—a seemingly insatiable appetite for cultural heritage both virtual and artifactual. Logic would tell us that increased demand for content would naturally increase demand for preservation of that content.

But logic would be wrong. Paradoxically, the proliferation of digital content, in high demand today, can make it harder to argue compellingly for preservation. Its sheer abundance and ubiquity makes digital content appear perdurable. And *mirabile dictu*, the Web provides masses of good-enough information to users without extracting a transaction fee. It is hard to see why we would need to start a public conversation about how to configure fair and equitable allocations of costs and benefits to ensuring long-term access to preserved content among societal sectors: whenever we go looking for information on a search engine, we find much more than we can use. There are no costs and everyone benefits, right?

Before we can even talk about who should bear the cost of long-term access to content of value—the ultimate public policy question—we should be able to argue compellingly why it matters. To date, we as a society have done poorly in making the case for public investment in preserving content. According to the evidence recently gathered in a national survey on the status of cultural content collections, our ability as a society to marshal resources to support preservation of our artifactual past is feeble (Heritage Preservation, 2005). Yes, government officials are known to wax eloquent on the heritage of our past and its value to present and future generations. But in reality it has fallen to the private sector to act as stewards of large parts of our recorded past, chiefly through philanthropy and an array of financial incentives featuring private ownership of publicly val-

ued culture and the intellectual property embodied in it. Libraries and archives—agents of the public trust—have benefited from a copyright regime that grants them exceptions for preservation of analog content. But the Section 108 exceptions to the copyright code are outmoded and fail to ensure preservation of many non-print analog-based cultural items, let alone digital content. It is entirely possible that current efforts to reform the law may be marginalized by the avoidance of copyright law altogether in favor of licensing access to privately held content.[4]

The report on the state of cultural collections, *A Public Trust at Risk: The Heritage Health Index Reports on the State of America's Collections* (Heritage Preservation, 2005), does not conclude that society is unwilling to pay for the stewardship of our collective cultural wealth. But it begs the questions of why the public trust is held in so little regard, and why its stewards are apparently unaccountable to the public, unable (or unwilling?) to rally resources for fulfilling the obligations they accepted when they accepted that trust.[5] That said, the systemic destabilization caused by our newest information technologies gives us an unprecedented opportunity to try to get it right again. The preservation community has an enviable chance to enlighten our society about the vital interests we collectively have in access to our cultural heritage, and to support the professionals that ensure that it is authentic, reliable, and easily found for access and use.

Doing Preservation: Analog Versus Digital

Digital content necessitates fundamental change in the methodologies and practices of stewardship throughout content's entire life cycle—including, of course, preservation. The good news is that not only the heritage sector is at work on this challenge. Our national security, economic, political, biological, and social well-being is now so critically dependent on reliable information management systems that it is not only preservationists who are grappling with the need to retool their infrastructures. Sectors that are information-intensive, from the military and law enforcement to medicine, science, engineering, manufacturing, agriculture, transportation, business and finance, insurance, higher education—are there any that are not?—find they require fundamentally new kinds of infrastructure to manage and preserve vital digital records. Infrastructure retooling requires a clear strategy to implement. It entails dedicating existing resources to new kinds of activities, establishing a series of long-term investments, rethinking the use of human resources and retraining of staff, expensive set-up costs for technologies, and so forth. It also entails finding new means of support and reassessing existing allocations of costs and benefits. Above all, it demands vision and leadership.[6]

The fundamental conceptual change that digital communication technology brings to organizations and to individuals is the need to manage

abundance, not scarcity (Hilton, 2006). Among the challenges of managing the much-too-muchness we must adapt to are:

- ever-increasing dependence on fragile technologies;
- adjusting preservation strategies to the different time-scales of digital content; and
- disappearing barriers to content creation resulting in the elimination of scarcity as a benchmark of value.

The Challenge of Technology

The rapid obsolescence of software and hardware is well known by now. But solutions to ensuring integrity and authenticity of complex content through hardware and software upgrades have not yet emerged. For the time being, we seem destined for information technology regimes in a state Comrade Trotsky would have recognized as "permanent revolution." Yes, storage is getting cheaper and cheaper. But storage is not preservation, a not so subtle distinction we have not yet made widely understood to technologists and non-specialists. No one is yet promising us that preservation will get cheaper and cheaper.[7]

The Challenge of Time-Scales

Until and unless we are able to automate nearly all processes for preservation after the initial decisions about what to collect are made, collecting, stabilizing or normalizing, and preserving digital content will remain a dynamic process demanding frequent human interventions. In contrast, analog preservation strategies can rely on some passive, very low-touch techniques, often as little as providing good housing. Though such passivity is seldom the best approach, it proves to be surprisingly forgiving at times. A remarkable amount of paper–based matter has survived years of deferred maintenance and outright neglect. But how would digital content fare under such a regime? Preservation of complex multimedia content—often the flavor primary sources of the day come in—requires more up-front planning, more active management, and possibly greater short- and long-term investments. The options for retroactive actions, including retrospective collecting, are few. So we need to act now, to expend resources now, to collect now. The reflex we have inherited from artifactual preservation practice is to justify present action against some future value. But how do we get a fix on the future value of digital content, especially in the cultural content area that fills the pages of the open Web?

The Challenge of Lower Barriers to Creation

Getting a fix on future value of digital content is hard in part because the technologies that make it so easy to consume digital content also make it dismayingly easy to create it, to distribute it widely, to appropriate con-

tent, change it, and create more, replicating this cycle and accelerating the growth of content. This can be done easily, promiscuously, and without barriers between creation and distribution. This may be excellent for us as citizens, and as consumers. It may be a boon to cultural heritage to give voice to many members of a culture. But of course for collectors, and especially for preservationists, it makes the business of selecting, collecting, preserving, and making available content rather more of a challenge. One answer to this challenge is not to choose at all but simply collect everything we possibly can, a strategy I shall return to.

Finally, there is the conceptual challenge of articulating intrinsic values inherent in content that is abundant rather than scarce. From a preservation point of view, it may be enough to say that certain attributes of cultural heritage objects that are benchmarks of value—rarity or uniqueness, antiquity, a physical object's aesthetic appeal and ability to be a repository of affect and association—are gone or at least marginalized. With artifacts, we can use the happy coincidence of uniqueness and physical instantiation to declare a thing that is scarce to be valuable simply because it is scarce. ("A rare book is valuable because it is rare.") Through reformatting, an incunable can have its intellectual content reprinted and physical attributes exquisitely represented in high-resolution scans; it can end up being more accessible for examination and at levels not attainable by the naked eye. But of course the imprint itself will continue to have value simply as a rare historical object. Its rarity has fixable market value as well. But we have no intrinsic scarcity in digital content, only the faux scarcity created by restrictions to access, usually designed to create market value and often resulting in decreased usefulness.

As frustrating as it may be, we cannot avoid trying in the present to determine the future values in digital content. Even if we decide not to make decisions about what to collect based on value and decide instead simply to collect as much as we can—an interesting thought experiment—we would still need to consider the values that inhere in the content. We need to identify value in order to know how to engineer the processes that will preserve content, that is, how precisely to ensure against loss of that value over time. So let us consider some aspects of value in content, before turning to the matter of value in preservation itself.

A Few Theories About the Value of Content

For the sake of brevity, I will forego discussion of the aesthetic, documentary, evidential, forensic, and economic or market values of content. These are values that are well known and understood by librarians and archivists and I have nothing original to contribute to their understanding in the current context. I will focus instead on aspects of content value that I believe behave differently in the digital realm or otherwise are not

obvious. The factors we must consider in order to get a deeper sense of just how crucial preservation is include:

- the usefulness of content (utilitarian value);
- the value of content as constitutive of our human nature, including its importance for ensuring biological homeostasis (species value) and its ability to give pleasure (hedonic value); and
- the value for re-use (secondary value).

The Value of Content Lies in Its Use

By virtue of its immateriality, digital content has the potential to be so much more useful to people than analog content. It is not embedded or exclusively instantiated in a unitary physical object that limits access to that content to one point in time and space. In other words, digital content is useful first and foremost because it is easy to use—easy to gain access to, computer-accessible anywhere, at any time, available both through push and through pull on smaller and evermore portable devices, and available to many people simultaneously. Digital access to content increases the autonomy of individuals as information-seeking creatures. Autonomy is a civic virtue highly prized by our society.

In the economics of information, use creates value. The more one uses content, the more valuable it becomes. An obvious corollary is that the selection of an item for use actually creates value. (This does not always equate to market value, but sometimes it does.) This is in fact the principle that underlies the page-ranking system used by Google, to take an obvious example. It is important to remember that a page rank is supposed to mean "people found this most useful," not "people found this to be better and of higher quality than everything else." Page rank expresses objective, that is, measurable, value. It does not exclude subjective value, but the search does not explicitly factor that in.

The usefulness of content is an important concept to grasp in thinking about societal value, and it is one that may strike some in curatorial and preservation communities as wrong-headed. To take an example: in the usual calculation of special collection libraries and archives, use is often seen to *demonstrate* or *validate* the value of something, but it is not credited with *creating* the value. That is, just because something is not used does not mean that it is not valuable. This understanding would argue for the notion that value is essentially intrinsic and cannot be conditioned by external factors such as use or even the perception of value. Perception of value is subjective and is conditioned by time and place. Some things that are currently prized by researchers and collectors will be neglected inside of a decade or two. The instances of collections languishing in book stacks untouched for decades and then coming into demand are legion. The increased use of cookbook collections, runs of Penny Dreadfuls, and other

collections perceived to be expressions of popular culture provide contemporary instances. Did years of nonuse actually decrease their value, or was that value simply unpotentiated?

In terms of assessing value, neither of those two questions address whether or not use itself increases value. I would argue that their current demand has, in fact, measurably increased their value. Because such nineteenth-century print collections as Penny Dreadfuls and housekeeping manuals are being used more, they demand more time of staff; because they are more at risk of being harmed (intentionally or not), they demand more staff and monetary resources to ensure their physical safety and usefulness. Items will often take a trip to the preservation department; they will receive more attention to their security in the reading room; and then of course, for the final "value-add," they will spend time in the scanning department to be reborn digitally. All the while, they "gain eyeballs," and not coincidentally, chances are they also start to gain market value. In many cases, library staff will track value of like items on eBay to gauge exactly how much the demand for such items increases. While it is critically important to hold firm the idea of the potential value of any given item of preserved content, it is equally important to acknowledge that it is the realization of that potential through use that is the goal of all collecting, curating, and preserving expenditures. The only possible reason that someone could wish for preserved content not to be fully used is because it puts the content at risk of losing its value through deterioration, defacement, theft, or other forms of degradation.

Significantly, these are not normal risks to digital content. Indeed, it is likely that, given the way preservation operates on digital content, the more digital content is used, the likelier it is to be preserved. And the more accessible content is—the more open the stacks, so to speak—the more likely the content will be found, tagged, cited, annotated, marked up, page ranked, recommended, and so on. It may or may not be true that information wants to be free, as people are fond of saying. But it would be true, if content were a species, that it would want to be open, because it would want to be found, to be used, to be replicated and preserved by staying in circulation—in short, to be fruitful and multiply.

The Species Value of Content

But what about digital content that is not frequently used—is it destined to die off? The fact that it might well would argue for collecting in the present as much digital content as we can possibly afford, even if we do not know what value it may have now, or in the future. Even if we can afford to do only "physical preservation" (that is, of the bits) and not be able to afford—or simply do not know how to do—"logical preservation" (maintain its renderability into something comprehensible to humans or machines), capturing content when you can is the safest strategy. Not only

is it a good bet to assume that technological innovation will continue and afford us the ability to decode the data later, but we should also assume that entirely new kinds of software may, through new types of data mining, find entirely new types of valuable information and expressiveness in preserved bits. To take yet another obvious example, think of the additional information that geographic information system (GIS) functionalities have given to existing data about where people live and where they work (demographic evidence of population shifts); where diseases cluster in a given population (epidemiological information about disease vectors); where people are registered to vote and in what number they turn out to vote (political information about civic engagement); what they purchase (market information about distribution patterns of consumption); and so on.

In the natural world, we can now track which particulate matter is found where in the atmosphere, what altitude certain plant species were found twenty years ago and where they are now (atmospheric and biological evidence of global warming); where tectonic plates underlying the Indonesian Archipelago were in November 2004 and where they were in January 2005 (seismic information); and in general, any combination of spatial and temporal data. Indeed, given the importance of tracking change over time in all four dimensions and the expanding functionalities of GIS, virtually no content about the natural world, including our impact on that world and our cultures as sites of biological phenomena (what I would call content with "species value"), is not worth capturing and keeping for some future use. In such a scenario, the decisions made about preservation do not center around *what* to preserve, but *how* to preserve, assessing types of value and using a cost-benefit analysis or other techniques to determine at what level of physical and logical preservation, what kinds of resources should be expended at which points in the expected life cycle of the content, and so forth.

In the case of much scientific and engineering data, it appears fairly straightforward to understand future uses of data, as their primary value for science and engineering rarely expires.[8] These domains of knowledge rely on the accumulation of data as well as of knowledge. Data, independent of knowledge, will retain their value long after a so-called paradigm shift in explanatory models renders that knowledge model ("paradigm") of historical interest only. We already know that we will need specific kinds of information about the disposal of toxic waste; about the power grid; about the load capacities of rivers, levees, and dams; about genomic, proteomic, and pharmacological molecules; about geological and seismic activities, and on and on. Here we must think about risk management strategies for preservation systems, as we should move away from mooting what we can afford to keep and focus rather on what, if anything, we can afford to lose. There are nascent efforts underway to further understand the

need for longitudinal data sampling and retention strategies, for database preservation, and for the data curation that these will demand.[9]

A Word on Behalf of Pleasure

What about the needs of cultural heritage content as distinct from scientific data? As readers will have noticed by now, I have not been making very fine or consistent distinctions between artistic, personal, scientific, geospatial, or other types of content. It is difficult to draw bright lines between the kinds of information that falls into the cultural heritage category and that which falls outside because the same content can have different significance in different contexts. For those who study humans and the products of their cultures, nearly all recorded content has value as the object of cultural study. Even content that is created with strictly scientific purposes can be prized for its historical, documentary, and aesthetic values. We know from the numerous natural history tomes that are now in rare book collections, such as illustrated botanicals, that scientific data having value for scientists for its record of past botanical species distribution, for example, take on a life of their own as precious commodities and aesthetic delights. Contemporary scientific data can have an equally aesthetic and hedonic value for the general public. Yes, the astronomical data that stream back from the Hubble telescope are not gripping to the average Web surfer when viewed as computer code. But NASA performs a kind of optical wizardry on those deep-space data to render them into mesmerizing, consciousness-altering images and posts them on their Web site in order to make their findings more accessible to the public and—lo and behold—people are affected intellectually and emotionally. Content has the power to provoke emotions such as awe and curiosity, both of which are pleasurable sensations. The information embedded in expressive content (which I consider those NASA images to be) is conditioned by the affect that it invokes and gains power thereby.

To the extent that any content performs cultural functions—recording experience, shaping perceptions of our world, adding or subtracting meaning, providing pleasure or inflicting pain—it should be construed as cultural. Cultural content plays a critical role in the production of meaning, and to the extent that it is used and shared, it also conduces to the development of empathy and social cohesion. It is in this role, among others, that cultural content is a pure public good, something that can be used by many yet never used up. Indeed, as I have been arguing, the more extensive the use by the greatest number of people, the more value it accrues and the greater the public good. That cultural content can serve many functions simultaneously simply indicates its value vector is manifold.

Salient differences between how content functions in a scientific context versus a cultural context have to do with different ways that the primary and secondary values of content operate. And these distinctions de-

rive in turn from the ways that content value is conditioned by different user communities.

Content has Primary and Secondary Values

Any given value in content is identified in the first instance by its creator or user. For what purpose was this item created, and who found that purpose to have value to them? Why was this commercial film made? To entertain and, if it was successful in doing so, it would make money. Why was this personal Web site created? To make public something the creator wished to communicate, and if it "got eyeballs," it succeeded in communicating that and possibly being linked to and getting more eyeballs. Why was this business record created? To comply with a regulation to do so and, if successful, to provide auditors what they required. There is little room for confusion about what the primary value might be in any given instance, because the intention of the creator is seldom hard to determine.[10]

The secondary value in content is a use above and beyond the use intended upon creation. It is the reason for re-use, and thus is the value that absorbs the attention of librarians, archivists, collectors, and connoisseurs. In some sense, it is only in the secondary value that preservation comes into play, in that it implies an investment has been made, or must be made, in order to make content available for use in a different context than its intended, primary use.[11] Usually, some effort is made to identify the potential secondary values of content to justify allocating preservation resources to its care. The value of preservation in this scenario is the value of using information at some time after the end of its normal accessibility for primary uses. In many cases, identification of secondary value is fairly straightforward, as in the case of the personal papers of important people, television news broadcasts, the recordings of famous performers, and so on.

But thinking that secondary value is always obvious has created regrettable lacunae in our cultural record. My favorite example—well known, vivid, and still painful—is the loss of close to 80 percent of early moving image materials. The primary use of cinema for decades was as commercial entertainment, as "product." Not enough people understood its secondary value as historical testimony and cultural heritage to make preserving it a priority, given that it is difficult to preserve. (Not only difficult: nitrate film is literally dangerous to keep around.) There is something very suggestive about the notion of preserved content—such as silent films on nitrate—as a form of stored energy, with the potential to come back into life to serve any number of unknown, perhaps unknowable, purposes. The denser the information storage medium, the more possible value it has recorded on it. Image technologies are incredibly dense information carriers and can convey so much information at once. In a melodrama from 1906 we can see the way people dressed, walked,

held their hands, and made gestures; or the shape of the landscape 100 years ago; what faces looked like before contemporary dentistry, universal inoculation against small pox, or modern nutrition. Digital content may in theory require significant upfront expenditure of today's resources to maintain that information potential over time. But it also has the promise of being more useful to us, because of its multimedia capabilities to record the world around us.

In many ways, the analogy with nitrate film is instructive by being misleading as well as suggestive. For our early film heritage, no matter how valuable it may or may not have seemed at the time or how difficult to preserve, still constitutes a very scarce form of content. You could have saved all of it and still have less than the content you can see on YouTube .com in five years—or is it one year?

The Value of Content Is Conditional

It stands to reason that if use of content creates value, then users are the agents of valuation. Yes, that sounds like the Web: the reputation systems that characterize Google, Amazon, eBay, MySpace, YouTube, Flickr, and so on are all based on user valuation of content. And yes, that is why they are wildly successful in providing value to users. But that principle is really old hat in the content world. Users have always been arbiters of content value and if the users are experts, of quality as well. In some contexts, we see the user-as-auditor who decides the value: does the record meet legally mandated standards of accountability. Or the user-as-consumer: what will I pay to acquire rights to use that content. Or the user-as-researcher: does that content meet my information needs. In all cases it is the expert user who determines the value of content for a specific purpose. The Internet has just changed who gets to be credited with expertise and how.

As a rule, societies are prejudiced in favor of experts, and for very good reasons. As societies become more complex, they develop prodigious amounts of expert knowledge that afford certain advantages for adaptation to the environment. This expert knowledge must be carefully stewarded in order for the next generation to use it. After a while, a society comes to be dependent on such knowledge. Scholarly experts in general are ideal arbiters of the value of academic content. Scientists are better at assessing the value of scientific data in their field than an untrained user, and so on. The openness of the Web, combined with Wiki technologies and others, have given rise to new models of expert contribution, editorial and peer review, and ways of making content of known value available. Here again we can point to certain advantages that digital information has over analog: it lowers the human overhead in assuring the quality and accessibility of information. The fact that in a large number of cases—certainly more than educators like—users will opt for good-enough informa-

tion rather than the best possible does not change that. Rather, it testifies to the judgments users make about how they wish to use their time.

The Value of Preservation

What does all this have to do with the value of preservation? Perhaps it is enough to point to the sum total of content's value to justify all our demands for preservation support. Utilitarian arguments tend to be persuasive to Americans. But I do not think that is an effective or reliable strategy to fund so essential an activity as preservation at the scale which it must now attain. Preservation may be necessary, but it is not necessarily instinctual. After all, humans are designed to prefer instant gratification over delayed gratification. Indeed, the implication of the common phrase "a bird in the hand is worth two in the bush" is that we are programmed to prefer short-term rewards over long. This is what economists have maintained for years and scientists more recently are validating, some proposing this reward preference as an evolutionarily advantageous strategy for survival that has been naturally selected. (The bird, after all, is a meal.) I also think that while some content owners are prepared to pay to secure the primary value of content, they are by and large not likely to spend sums to create the possibility for others to mine secondary values. Those charged with securing the secondary value of content, usually memory institutions, lack sufficient resources to do so, a finding that has been sadly documented by the Heritage Health Index. That is why cultural content needs to have greater protection from risk of loss than our society currently musters. And that is why we need to articulate specific values in the business of preservation itself.

So let me extrapolate a few of the distinct affordances created by preservation that have been identified so far in this paper:

- Preservation is the cost of access
- Preservation is insurance against loss of value
- Preservation protects against loss of business continuity in the event of disruptions and catastrophes
- Preservation protects our critical information-based dependencies
- Preservation adds value to content by maximizing its potential for reuse

These are all concrete benefits one could factor into a cost/benefit analysis or could cite to persuade decision makers to spend money on a risk management and mitigation strategies, on disaster preparedness and response plans, on insurance against loss, and to ensure business continuity.

The ultimate societal benefit of preservation is, of course, to ensure the well-being of the population and the survival of our society, and indeed, our species. Given that information is a constitutive force in society,

all aspects of its integrity, completeness, authenticity, and accessibility are profoundly important. And all choices about whether we do or do not decide to preserve historical records and cultural expressiveness from times past are themselves constitutive choices (Braman, 2006; Starr, 2004, pp. 4–5). As Russians are fond of reminding themselves, "The future is certain, it is the past that is unpredictable." A people who do not own and control their own cultural heritage are a people who can be held captive by false histories, fabrications, and lies. The genius of totalitarian societies is that the need for brute force to make subjects out of citizens is really quite modest. If the government controls what people know about their past and their present, they limit the scope of their imaginations and can control their expectations for the future. We may hold dear the notion that because we are not a totalitarian country, we are not at risk of developing false memories, fabrications, and blank spots in our past. But that complacency is dangerous. Memory can play tricks on all of us.

In truth we need both historical knowledge (knowledge about the past) and memory (knowledge from the past) to be rich, complete, authentic, and reliable, for the social cohesion and moral life of a community (Margalit, 2004, p. 114). The need for preservation in the service of historical knowledge is well understood, in particular by librarians and archivists. But our understanding of memory—how it works biologically and the function it plays in our destiny as individuals and as a species—is a rather more novel investigation. We are beginning to understand that memories are emergent neural phenomena, literally the remembering of discrete packets of information and affect. In order to have functional memory, the information packets must be intact and the affect unimpaired, able to provide accurate clues about the importance, priority, and purpose of the information held in memory. And the emergent memory must find the right context in a web of other memories that can situate its meaning in the present. [12] Senile dementia, amnesia, and Alzheimer's are examples of what happens when the remembering system as a whole does not work because constituent parts begin to fail. In the case of Alzheimer's, we can point to physical holes in parts of the brain that may account for some of the failure to remember information and attach the appropriate affect to it. In the historical fabric of any culture, there will be such holes as well, and there we find confusion, confabulation, loss of identity and of purpose. What is uniquely important about cultural content, as opposed to factual information and knowledge about our world, is the moral value of having access to knowledge from the past, in its authentic and unmediated form. While I will not get into the "notorious intellectual quagmire" that is speculation about the evolutionary basis of morality, I will simply assert that ethics count and constitute another reason preservation has societal value (Coyne, 2006, p. 983).

Public Policy Implications of the Value of Preservation

In order to better secure a whole, accurate, and useful shared record of the past, we need to make several things happen:

- Make it easier and cheaper to preserve content
- Provide incentives and rewards for individuals and organizations to preserve
- Protect the public interest in privately held content

Make It Easier and Cheaper to Preserve Content

Preservation of both artifactual and digital content could be less expensive than it is if it were properly engineered. For both analog and digital content, it is better to invest in prevention and measures taken upstream than down. For artifacts, that would mean better storage environments, with disaster preparedness and recovery plans in place, stabilization of content through such means as deacidification, rehousing, transferring to stable media, and so on. Such an approach is appropriate for born-digital content as well: we should be aiming to make content "born-archival," in relatively stable or ubiquitous formats, replete with automatically generated metadata, regularly backed up off-site, and residing on systems that are geared to the "permanent revolution" of technology changes and upgrades. The preservation communities that support different media could develop a technical research and development agenda whose specific aim would be to test promising preservation-friendly technologies and move proofs of concept into production mode.

Provide Incentives and Rewards to Preserve

In order to ramp-up preservation activities at a sufficiently large scale, we shall need to mobilize both organizations that produce content and individuals who care about it. How do we provide incentives for creators of digital content to make the content born-archival? This is clearly an important area for investigation, and the economics of archiving is a woefully underdeveloped field. Given the values and characteristics of content that I suggest truly matter when thinking of society's willingness to pay—conditioned by time and place, contextualized by use, having community-defined values that can be widely and simultaneously shared yet seldom at risk of depletion, and essential for the well-being of society—it seems that we would have a good deal to learn from environmental economics. They, too, look at issues of balancing short-term and long-term time horizons, providing incentives to private entities to protect the public interest, developing tax and other financial policies that attempt to monetize the benefits of doing so, and so on.

It is not enough to insist on the value of preservation without working

to create real-world strategies to promote and protect it. While it is desirable, of course, for governmental entities and prestigious international bodies to declare the value of preserving content, it becomes counter-productive when they do so without making provision for financial support and compliance requirements. Digital preservation does not need to join artifactual preservation as a worthy but bankrupt ideal, yet another unfunded mandate. The information landscape is littered with requirements to preserve or deposit data in trustworthy archives, yet there is little funding earmarked to do so and few unpleasant consequences when directives are ignored.[13] Advocating that private companies which create content start to be responsible for preserving it for the collective good will fall on deaf ears. For them as well it would become an unfunded mandate.

Protect the Public Interest in Privately Held Content

It becomes readily apparent that cultural content that does not receive resource investments that maximize potential for reuse, such as metadata creation and tagging, open distribution, spidering, and redundant downloading, curation, and asset management, has very little chance of surviving for long after it has lost its primary market. Content that is actively used for its primary value probably faces little risk of loss; and the growth of the long-tail market for some creative content assures a longer life span for a significant portion of books, movie, and music that might otherwise become inaccessible. But let us imagine a situation in which some segment of society wishes to have access to some content that is privately held and not actively used, or is negligently stewarded—the assets of a defunct recording studio, say, or the papers of an influential journalist. Is that public entitled to have access to that content for the sake of enjoying its beauty, taking pleasure from it, or using it for educational purposes? We may not have the right to own it, but do we not have some interest in access to it at some point in time? What protection do we have against the loss of such content?[14]

Not much, it turns out. The "intangible elements of our cultural heritage such as arts, skills, folklife, and folkways" are protected by the National Historic Preservation Act Amendments of 1980 (Public Law 96-515) (American Folklife Center, 1983).[15] In 1988 the National Film Preservation Act was enacted (and has been subsequently reauthorized, most recently in 2005), to identify and preserve moving images of national significance.[16] The National Recording Preservation Act of 2000 likewise is designed "to maintain and preserve sound recordings that are culturally, historically, or aesthetically significant."[17] These are excellent legislative recognitions of the importance of cultural heritage to the country. But they are hampered by having to work within a copyright regime that is counter-productive to their aims. Congress long ago bought the argument that extending copyright monopoly over content such as recorded

sound, image, and text would provide owners an incentive to preserve it and make it available to the public for prices that the market would bear. This has not turned out to be an entirely successful strategy.[18] It would be preferable to let such material that is either neglected by its owners or orphaned into the hands of the public and let it find those who are willing to adopt it, take care of it, and protect its value. This is a role that collectors have played for centuries, and we do not even need to be creative about inventing incentives to encourage them to do so. If we were to open up access to this kind of content on the Web, it would be even easier to attract users who would describe it, tag it, recommend it, and increase its use and its value. I am not arguing for requiring private owners to preserve content at their own expense for the public good. I wish they would, but to mandate this amounts to an unfunded mandate and is not likely to achieve any of the desired effects. I do think that we should begin a national conversation about how to enable content owners to continue enjoying the benefits of owning content, especially its primary value benefits (which is usually making it available in the market place), while ensuring adequate stewardship of it for the collective good. We should condition the privileges of ownership by the obligations of ensuring that potential for secondary use—cultural reuse in particular—not be nullified. There are, of course, many possible ways to make this happen. (This is what one possible function of a dark or dimly lit archive would serve, for example.)

What about content that is already publicly available? There is something extremely appealing about the model we see with certain kinds of content available on the open Web, much of it of cultural significance. It is the kind that benefits from this wonderful economic model of "give it away, build reputation, charge for value-added goods or services" such as ad space, higher resolution version of the content, a hard copy. How can we make this work for preservation of cultural content?

Finally, let me argue for additional incentives and rewards for those institutions that act as surrogates for the public—the libraries and archives, museums and herbaria, historical societies, data repositories, and other stewardship organizations that act on behalf of the public to provide preservation of socially significant content. I believe that they should continue to benefit from whatever legislative and funding considerations they need to discharge the public trust. I would also argue that they need to be more publicly accountable for what they do, or do not do, so that they may demonstrate that the trust the public places in them is well tended. We have a right to expect that they will also be stronger and more effective advocates for preservation than they have been and be able to articulate the societal benefits of stewardship in the language that others speak. Several years ago I was admonished by a scientist who was actively engaged in establishing data curation protocols—preservation requirements—for his field

that librarians should never use the word "preservation" when what they really mean is access. Yes, he said in effect, it is access to noncurrent data, and there are many processes one needs to undertake in order to afford the use of noncurrent data. And yes, librarians may call that preservation. But one must always speak of processes, especially complicated and expensive processes, in terms of their end result. Then people can follow the conversation with sympathy and interest.

I cite this anecdote in part to suggest how others think technical issues of data curation and preservation should be discussed in mixed company. But I also wish to applaud the sheer common sense that tells us that when we enter the arena of public policy—as I believe we must—then it is the benefits of preservation that are most vividly articulated. I do not underestimate how difficult that can be. Nor do I think that the societal benefits of preservation that I have sketched here are a complete list of all that our society has to gain from having access to knowledge from the past. Indeed, these are only the beginning. I believe that collaborations with cultural anthropologists and neuroscientists, economists and engineers, ethicists and historians would deepen our understanding of these benefits, help to develop compelling arguments for their support, and suggest real-world solutions to making preservation cheaper, easier, and more rewarding.

Acknowledgments

I wish to thank Samuel Brylawski, Paul Courant, Beth Dulabahn, David Rumsey, and Steven Wheatley, from whom, through a series of conversations, which were often about other topics, I have learned much about preservation's value to society. While they may be presumed to share an interest in the societal value of preservation, they may not be presumed to share some—or any—of the views I express here.

Notes

1. To take the case of imprints: we understand the basic photochemical and thermal processes that lead to the degradation of paper. We know how to retard and even arrest that degradation through such means as deacidifying paper, providing ideal storage environments, and creating use surrogates to reduce handling originals. We know how to achieve economies of scale for all these measures as well, even if we have not always implemented them.
2. I wish to avoid the confusing term *intangible cultural asset*, as often the content this refers to is in fact recorded onto tangible media. In this essay I refer to content as the products of intellectual and artistic activity that are textual, visual, aural, and numeric, information-based, and recorded on a medium, either analog or digital. So I exclude the built and unbuilt environments and cultural products that are not recorded onto media. Content is often subject to copyright regulations. In some instances I refer to data gathered by humans, which is, of course, not subject to copyright.
3. Six research libraries have embarked on individual partnerships with Google to digitize large parts of their book collections: the Bodleian Library, the New York Public Library, the University of Michigan Library, Harvard College Library, Stanford University Libraries, and the University of California (UC) library system. As the press release from UC clearly

shows, the libraries themselves are committing significant resources to the projechttp://www.universityofcalifornia.edu/news/2006/aug09.html

4. See activities of the Section 108 Working Group, sponsored by the U. S. Copyright Office, at http://www.loc.gov/section108/; and reports by June Besek on copyright law as it relates to recorded sound preservation and to digital archiving. (Besek, 2003, 2005).
5. The report (Heritage Preservation, 2005) cites that fully 80 percent of collecting institutions "do not have an emergency or disaster plan that includes collections, with staff trained to carry it out. . . . 70% have no current assessment of the condition of their collections. . . . only 2% of the total annual budget of U.S. collecting institutions was spent on preservation in the last fiscal year" (p. 2). Despite the latter, fully 40 percent report having not even tried to raise funds for preservation of collections they hold "in the public trust" (p. 77).
6. These are all themes raised in reports on developing cyberinfrastructure capacities to support education and research, issued by the American Council of Learned Societies and the National Science Foundation (ACLS, 2006; NSF, 2005, 2006).
7. Although Chris Rusbridge (2006) argues to the contrary.
8. To illustrate the point, let me take almost at random a recent article from a scientific journal. "The Continuous Plankton Recorder survey, started in 1946, maintains population records on these key microorganisms at the bottom of the ocean's food chain" (Kintisch, 2006, p. 778). According to Patrick Halpin, a researcher who has used the analysis of fifty years of data from this survey, "'historic records are so valuable when you start thinking about [climate] change. There are so many things like that we wished we had done'" (p. 778). Not only are longitudinal datasets very important, but the ability to combine them with others is increasingly recognized as fruitful. "As their analyses are getting more sophisticated, marine ecologists considering the impact of climate change are seeking more interdisciplinary approaches and combining different kinds of data more extensively" (p. 778).
9. See NSF (2005, 2006).
10. Except, of course, in those cases in which the context of creation and use are wholly obscure, or the content unintelligible, factors not uncommon in cultural content.
11. Work done to extend primary value could be called preservation as well, but it is more often called information management, digital asset management, or data curation.
12. A rich and accessible description of the role of emotion in information processing and memory formation, incorporating a range of experimental findings, can be found in D'Amasio (2000).
13. One example is the recent EU Recommendation (24 August 2006), building on a previous EU Resolution (25 June 2006), that calls for member states to digitize content and preserve born-digital; it cites the existence of legal mandates already in place to deposit digital content into repositories. http://europa.eu.int/information_society/actvities/digital_libraries/index.htm.
14. This is a topic that Joseph L. Sax investigates in his book, *Playing Darts with Rembrandt* (1999). While he makes the case only in reference to "cultural treasures" such as famous works of art, I would argue that the case for cultural content is the same. Obviously, in light of the Sarbanes-Oxley regulation that is leading to large-scale destruction of corporate records, this area should come under consideration as well.
15. Carl Fleischhauer kindly provided me with this source.
16. For the full history of the Act, see http://www.loc.gov/film/filmabou.html.
17. From P. L. 106-474, available at http://www.loc.gov/rr/record/nrpb/nrpb-home.html.
18. A recent study by Tim Brooks (2005), for example, "finds that most U.S. historical sound recordings have become virtually inaccessible—available neither commercially nor in the public domain. According to the report, the rights to 84 percent of historically significant recordings made in the United States between 1890 and 1964 are still owned by someone and are therefore protected by law. For most pre-1972 recordings, protection comes in the form of state, not federal, law until 2067. Because recordings cannot be copied and distributed without permission of their rights holders, the only legal way to obtain a CD of a pre-1972 recording is through a reissue. Yet the study found that rights holders have reissued—or allowed others to reissue—on CD only 14 percent of the pre-1965 recordings they control. Thus, most historically important sound recordings are available for

hearing only through private collectors or at research libraries that collect our audio heritage and have the equipment to play obsolete, often frail recordings.

References

American Council of Learned Societies. (2006). *Our cultural commonwealth: The final report of the American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities & Social Sciences.* New York: American Council of Learned Societies. Retrieved January 16, 2007, from http://www.acls.org/cyberinfrastructure.

American Folklife Center. (1983). *Cultural conservation: The protection of cultural heritage in the United States.* Washington, DC: Library of Congress.

American Heritage Dictionary (4th ed.). (2000). Boston: Houghton Mifflin.

Besek, J. M. (2003). *Copyright issues relevant to the creation of a digital archive: A preliminary assessment* (CLIR Report 112). Washington, DC: Council on Library and Information Resources and the Library of Congress. Retrieved January 16, 2007, from http://www.clir.org/pubs/abstract/pub112abst.html.

Besek, J. M. (2005, January). *Copyright issues relevant for digital preservation and dissemination of pre-1972 commercial sound recordings by libraries and archives* (CLIR Report 135). Washington, DC: Council on Library and Information Resources and the Library of Congress. Retrieved January 16, 2007, from http://www.clir.org/pubs/abstract/pub135abst.html.

Braman, S. (2006). Tactical memory: The politics of openness in the construction of memory. *First Monday, 11*(7). Retrieved January 16, 2007, from http://firstmonday.org/issues/issue11_7/braman/index.html.

Brooks, T. (2005, August). *Survey of reissues of U.S. recordings* (CLIR Report 133). Washington, DC: Council on Library and Information Resources and the Library of Congress. Retrieved January 16, 2007, from http://www.clir.org/pubs/abstract/pub133abst.html.

Coyne, J. A. (2006, August 31). Selling Darwin: Does it matter whether evolution has any commercial applications? *Nature, 442,* 983–984.

D'Amasio, A. (2000). *The feeling of what happens: Body and emotion in the making of consciousness.* New York: Harcourt.

Heritage Preservation. (2005). *A public trust at risk: The Heritage Health Index report on the state of America's collections.* Washington, DC: Heritage Preservation.

Hilton, J. (2006, March/April). The future for higher education: Sunrise or perfect storm? *EDUCAUSE Review, (41)*2, 59–71. Retrieved January 16, 2007, from http://www.educause.edu/apps/er/erm06/erm062.asp.

Kintisch, E. (2006, August 11). As the seas warm. *Science, 313,* 776–779.

Lyman, P., & Varian, H. R. (2003). *How much information? 2003.* Retrieved January 16, 2007, from http://www2.sims.berkeley.edu/research/projects/how-much-info-2003.

Margalit, A. (2004). *The ethics of memory.* Cambridge, MA: Harvard University Press.

National Recording Preservation Act, 2 USC § 1701 (2000).

National Science Foundation. (2005). Revolutionizing science and engineering through cyberinfrastructure: Report of the Blue Ribbon Advisory Panel on Cyberinfrastructure. Washington, DC; National Science Board. Retrieved January 16, 2007, from http://www.nsf.gov/od/oci/reports/toc.jsp.

National Science Foundation. (2006). NSF's cyberinfrastructure vision for 21st century discovery (v. 5.0). Washington, DC: National Science Foundation.

Rusbridge, C. (2006). Excuse Me . . . Some Digital Fallacies? *Ariadne, 46.* Retrieved January 16, 2007, from http://www.ariadne.ac.uk/issue46/rusbridge.

Sax, J. L. (1999). *Playing darts with Rembrandt: Public and private rights in cultural treasures.* Ann Arbor: University of Michigan Press.

Starr, P. (2004). *The creation of the media: Political origins of modern communications.* New York: Basic Books.

Abby Smith is a historian and consulting analyst focusing on the creation, preservation, and use of the cultural record in a variety of media. She works with the Library of Congress's National Digital Information Infrastructure and Preservation Program in development of its national strategy to identify, collect, and preserve digital content

of long-term value. Director of programs at the Council on Library and Information Resources until 2005, she has also worked at the Library of Congress managing programs relating to preservation of and access to cultural heritage collections. She is currently working with several universities on identifying content of long-term value, understanding various risk factors to its persistence, and analyzing organizational strategies for its long-term access.

She holds a doctoral degree in history from Harvard University and has taught at Harvard and Johns Hopkins Universities. Her recent publications include: *Access in the Future Tense*; *New-Model Scholarship: How Will It Survive?; Strategies for Building Digitized Collections*; *The Evidence in Hand: Report of the Task Force on the Artifact in Library Collections*; and *Authenticity in the Digital Environment.*